BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 7: e111662
doi: 10.3897/biss.7.111662

OPEN ACCESS

Conference Abstract

Name IDs and Name Matching for Catalogue of Life: Existing Services and Prospects

Olaf Bánki[‡,§], Markus Döring[|,§], Thomas S. Jeppesen[|]

‡ Naturalis Biodiversity Center, Leiden, Netherlands
§ Catalogue of Life / Species 2000, Weesp, Netherlands
| Global Biodiversity Information Facility, Copenhagen, Denmark

Corresponding author: Olaf Bánki (olaf.banki@sp2000.org)

Received: 25 Aug 2023 | Published: 28 Aug 2023

Citation: Bánki O, Döring M, Jeppesen TS (2023) Name IDs and Name Matching for Catalogue of Life: Existing Services and Prospects. Biodiversity Information Science and Standards 7: e111662.
https://doi.org/10.3897/biss.7.111662

Abstract

ChecklistBank, developed by Catalogue of Life (COL) and the Global Biodiversity Information Facility (GBIF), is a publishing platform and open data repository focused on taxonomic and nomenclatural data sets (checklists). It contains close to 50K datasets, mostly originating from digitised peer reviewed scientific articles mediated by Plazi, amongst others. The COL Checklist (Bánki et al. 2023) is assembled out of a selection of the data sources in ChecklistBank. The Catalogue of Life Checklist is issued with name usage identifiers, as well as a digital object identifier for the Checklist version (with an associated dataset key). The more than 160 data sources that make up the COL Checklist are also issued with digital object identifiers as well as a data set key. The combination of a name usage identifier and the data set key allows for the tracking of names between the various COL Checklist versions. ChecklistBank is built in an open API. It supports data sharing through various exchange formats of the Darwin Core (Darwin Core Task Group 2009) data standard (e.g., Darwin Core-Archives (GBIF 2021) and ColDP), and provides several download and name matching options.

The Transforming European Taxonomy through Training, Research, and Innovations (TETTRIs) European Union funded project will contribute to a couple of improvements to ChecklistBank. In the context of the TETTRIs project, a new name usage (i.e., taxon or synonym) matching service against any dataset in ChecklistBank, not just the COL

© Bánki O et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Checklist, was developed. A single name matching service takes query parameters for a single name and optionally its classification. The service allows for bulk matching of names against the ChecklistBank API. This contains the option of matching a classification in a CSV file. The bulk matching allows all names of an entire or a subtree of an existing ChecklistBank dataset to act as the source for names instead of the input matching a CSV file. The bulk matching services are asynchronous and notify a user by email when the results are ready to be downloaded in a CSV file.

Keywords

taxonomy, ChecklistBank, ColDP, data standards

Presenting author

Markus Döring

Funding program

The Transforming European Taxonomy through Training, Research, and Innovations (TETTRIs) is a EU funded project with grant number 101081903.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Bánki O, Roskov Y, Döring M, Ower G, Hernández Robles DR, Plata Corredor CA, Stjernegaard Jeppesen T, Örn A, Vandepitte L, Hobern D, Schalk P, DeWalt RE, Keping M, Miller J, Orrell T, et al. (2023) Catalogue of Life Checklist (Annual Checklist 2023). Catalogueoflife.org. URL: https://doi.org/10.48580/dfsr
- Darwin Core Task Group (2009) Darwin Core. Biodiversity Information Standards (TDWG). URL: http://www.tdwg.org/standards/450
- GBIF (2021) Darwin Core Archives – How-to Guide, version 2.2. Copenhagen: GBIF Secretariat. URL: https://ipt.gbif.org/manual/en/ipt/2.5/dwca-guide